⊙ PⓇRTAL

USPTO

LDI tree: a hierarchical representation for image-based rendering

Full text   📄Pdf (1.09 MB)

Source   International Conference on Computer Graphics and Interactive Techniques archive
Proceedings of the 26th annual conference on Computer graphics and interactive techniques table of contents
Pages: 291 - 298
Year of Publication: 1999
ISBN:0-201-48560-5

Authors   Chun-Fa Chang University of North Carolina at Chapel Hill
Gary Bishop   University of North Carolina at Chapel Hill
Anselmo Lastra University of North Carolina at Chapel Hill

Sponsor   SIGGRAPH: ACM Special Interest Group on Computer Graphics and Interactive Techniques

Publisher   ACM Press/Addison-Wesley Publishing Co.   New York, NY, USA

Additional Information: references   citings   index terms   collaborative colleagues   peer to peer

Tools and Actions: Find similar Articles   Review this Article
Save this Article to a Binder   Display Formats: BibTex   EndNote   ACM Ref

DOI Bookmark: Use this link to bookmark this Article: http://doi.acm.org/10.1145/311535.311571
What is a DOI?

↑ REFERENCES

Note: OCR errors may be found in this Reference List extracted from the full text article. ACM has opted to expose the complete List rather than only correct and linked references.

1 C. H. Chien, Y. B. Sim and J. K. Aggarwal. Generation of Volume/Surface Octree from Range Data. The Computer Society Conference on Computer Vision and Pattern Recognition, pages 254-60, June 1988.

2 C. I. Connolly. Cumulative Generation of Octree Models from Range Data. Proceedings, Int'l Conf. Robotics, pages 25-32, March 1984.

3 Brian Curless , Marc Levoy, A volumetric method for building complex models from range images, Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, p.303-312, August 1996

4 Steven J. Gortler , Radek Grzeszczuk , Richard Szeliski , Michael F. Cohen, The lumigraph, Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, p.43-54, August 1996

5 Craig Kolb. Rayshade. http://www-graphics.stanford.edu/~cek/rayshade/.

6 David Laur , Pat Hanrahan, Hierarchical splatting: a progressive refinement algorithm for volume rendering, ACM SIGGRAPH Computer Graphics, v.25 n.4, p.285-288, July 1991

7 Marc Levoy , Pat Hanrahan, Light field rendering, Proceedings of the 23rd annual conference on

LDI Tree: A Hierarchical Representation for Image-Based Rendering

Chun-Fa Chang   Gary Bishop   Anselmo Lastra

University of North Carolina at Chapel Hill

ABSTRACT

Using multiple reference images in 3D image warping has been a challenging problem. Recently, the Layered Depth Image (LDI) was proposed by Shade et al. to merge multiple reference images under a single center of projection, while maintaining the simplicity of warping a single reference image. However it does not consider the issue of sampling rate.

We present the LDI tree, which combines a hierarchical space partitioning scheme with the concept of the LDI. It preserves the sampling rates of the reference images by adaptively selecting an LDI in the LDI tree for each pixel. While rendering from the LDI tree, we only have to traverse the LDI tree to the levels that are comparable to the sampling rate of the output image. We also present a progressive refinement feature and a "gap filling" algorithm implemented by pre-filtering the LDI tree.

We show that the amount of memory required has the same order of growth as the 2D reference images. This also bounds the complexity of rendering time to be less than directly rendering from all reference images.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation - Viewing Algorithms; I.3.6 [Computer Graphics] Methodology and Techniques - Graphics data structures and data types; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism.

Additional Keywords: image-based rendering, hierarchical representation

1. INTRODUCTION

The 3D Image warping algorithm [14] proposed by McMillan and Bishop uses regular single-layered depth images (which are called *reference images*) as the initial input. One of the major problems of 3D image warping is the disocclusion artifacts which are caused by the areas that are occluded in the original reference image but visible in the current view. Those artifacts appear as tears or gaps in the output image. In Mark's Post-Rendering Warping [11], the techniques of splatting and meshing are proposed to deal with the disocclusion artifacts. Both splatting and meshing are adequate for post-rendering warping in which the current view does not deviate much from the view of the reference image.

However, the fundamental problem of the disocclusion arti-

CB#3175 Sitterson Hall, Chapel Hill, NC 27599-3175, USA.
{chang, gb, lastra}@cs.unc.edu http://www.cs.unc.edu/~ibr

facts is that the information of the previously occluded area is missing in the reference image. By using multiple reference images taken from different viewpoints, the disocclusion artifacts can be reduced because an area that is not visible from one view may be visible from another. When multiple source images are available, we expect the disocclusion artifacts that occur while warping one reference image to be eliminated by one of the other reference images. However, combining multiple reference images and eliminating the redundant information is a non-trivial problem, as pointed out by McMillan in his discussion of inverse warping [15].

Recently, the Layered Depth Image (LDI) was proposed by Shade et al. [19] to merge many reference images under a single center of projection. It tackles the occlusion problems by keeping multiple depth pixels per pixel location, while still maintaining the simplicity of warping a single reference image. Its limitation is that the fixed resolution of the LDI may not provide an adequate sampling rate for every reference image. Figure 1 shows two examples of such situations. Assuming the two reference images have the same resolution as the LDI, the object covers more pixels in reference image 1 than it does in the LDI. Therefore the LDI has a lower sampling rate for the object than reference image 1. Similar analysis shows the LDI has a higher sampling rate than reference image 2. If we combine both reference images into the LDI and render the object from the center of projection of reference image 1, the insufficient sampling rate of the LDI will cause the object to look more blurry than it looks in reference image 1. When we render the object from the center of projection of reference image 2, the excessive sampling rate of the LDI might not hurt the quality of the output. However, processing more pixels than necessary slows down the rendering.

In this paper, we present the *LDI Tree*, which combines a hierarchical space partition scheme with the concept of the LDI. It preserves the sampling rate of the reference images by adaptively selecting an LDI in the LDI tree for each pixel. While rendering from the LDI tree, we only have to traverse the LDI tree to the levels that are comparable to the sampling rate of the output image. Because each LDI also contains pre-filtered results from its children LDIs, progressive refinement is easy to implement. The pre-filtering also enables a new "gap filling" algorithm to fill the disocclusion artifacts that cannot be resolved by any reference image.

The amount of memory required has the same order of growth as the 2D reference images. Therefore the LDI tree preserves an important feature that image-based rendering has over traditional polygon-based rendering: the cost is bounded by the complexity of the reference images, not by the complexity of the scene.

2. RELATED WORK

2.1. Inverse Warping

The image warping described in [14] is a forward warping process. The pixels of the reference images are traversed and warped to the output image in the order they appear in the reference images. Some pixels in the output image may receive more than
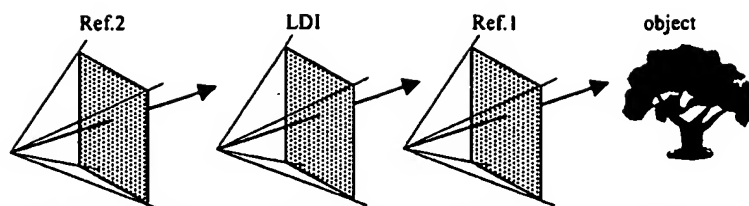
Figure 1: The LDI does not preserve the sampling rates of the reference images.

one warped pixel and some may receive none, which causes artifacts.

In [15], McMillan proposed an inverse warping algorithm. For each pixel in the output image, searches are performed in all reference images to find the pixels that could be warped to the specified location in the output image. Although epipolar geometry limits the search space to a one-dimensional line or curve in each reference image and a quadtree-based optimization has been proposed in [10], searching through all reference images is still time consuming.

2.2. Layered Depth Image

Another way to deal with the disocclusion artifacts of image warping is to use the Layered Depth Image (LDI)[19]. Given a set of reference images, one can create an LDI by warping all reference images to a carefully chosen camera setup (e.g. center of projection and view frustum) which is usually close to the camera of one of the reference images. When more than one pixel is warped to the same pixel location of the LDI, some of them may be occluded. Although the occluded pixels are not visible from the viewpoint of the LDI, they are not discarded. Instead, separate layers are created to store the occluded pixels. Those extra pixels are likely to reduce the disocclusion artifacts. However the fixed resolution of the LDI limits its use as discussed previously in section 1.

Lischinski and Rappoport used three parallel-projection LDIs to form a Layered Depth Cube [9]. Max's hierarchical rendering method [12] uses the Precomputed Multi-Layer Z-Buffers which are similar to the LDIs. It generates the LDIs from polygons and the hierarchy is built into the model.

2.3. Volumetric Methods

The LDI resembles volumetric representations. The main differences between an LDI-based representation and 3D volume data are discussed in [9].

Curless and Levoy presented a volumetric method to extract an isosurface from range images [3]. The goal of their work, however, was to build high-detail models made of triangles. The volume data used in that method is not hierarchical and it relies on a run-length encoding for space efficiency.

There has also been work related to octree generation from range images [1][2][8]. However the octree that is generated in those methods is used to encode the space occupancy information. Each octree cell represents either completely occupied or completely empty parts of the scene.

The multi-resolution volume representation in the Hierarchical Splatting work [6] by Laur and Hanrahan can be considered as a special case of the LDI tree in which the LDIs are of 1×1 resolution. It is however built from a fully expanded octree (which is called a pyramid in their paper). The octree to be traversed during the rendering is also predetermined and does not change with the viewpoint.

2.4. Image Caching for Rendering Polygonal Models

The image caching techniques of Shade et al. [18] and Schaufler et al. [17] use a hierarchical structure similar to the LDI tree. Each space partition has an imposter instead of an LDI. The imposter can be generated rapidly from the objects within the space partition by using hardware acceleration. However, the imposter has to be frequently regenerated whenever it is no longer suitable for the new view.

In contrast, the information stored in the LDI tree is valid at all times. By generating the LDI tree from the reference images instead of the objects within the space partitions, the LDI tree can be used for non-synthesized scenes as well.

3. LDI TREE

The LDI tree is an octree with an LDI attached to each octree cell (node). The octree is chosen for its simplicity but can be replaced by the other space partitioning schemes. Each octree cell also contains a bounding box and pointers to its eight children cells. The root of the octree contains the bounding box of the scene to be rendered[1]. The following is pseudo code representing the data structure:

```
LDI_tree_node =
Bounding_box[X..Z, Min..max]: array of
    real;
Children[0..7]: array of pointer to
    LDI_tree_node;
LDI: Layered_depth_image
```

All LDIs in the LDI tree have the same resolution, which can be set arbitrarily. The height (or number of levels) of the LDI tree will adapt to different choices of resolution. In general, a lower resolution results in more levels in the LDI tree. Ultimately, we can make the resolution of the LDI be 1×1 which makes the LDI tree resemble the volume data in the Hierarchical Splatting [6].

Note that each LDI in the LDI tree contains only the samples from objects within the bounding box of the cell. This is sometimes confusing because the LDI originally proposed by Shade et al. combines the samples from all reference images.

For simplicity, we use one face of the bounding box as the projection plane of the LDI. Orthographic projection is used and the projection direction is perpendicular to the projection plane.

An example of the LDI tree is shown in Figure 7 by viewing the bounding boxes from the top. The following sections discuss the details of constructing the LDI tree from multiple reference images and of rendering a new view from the LDI tree.

---

[1] For outdoor scenes, background textures can be added to the faces of the bounding box. The bounding box can be extended with little overhead if most of the space is empty.
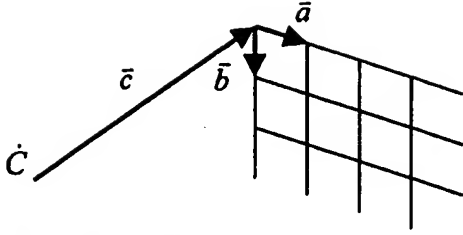
Figure 2: The camera model.

3.1. Constructing the LDI Tree from Multiple Reference Images

The LDI tree is constructed from reference images by warping each pixel of the reference images to the LDI of an octree cell, then filtering the affected LDI pixels to the LDIs of all ancestor cells in the octree.

In 3D image warping, each pixel of the reference images contains depth information which is either stored explicitly as a depth value or implicitly as a disparity value. This allows us to project the center of the pixel to a point in the space where the scene described by the reference images resides.

We observed that the sampling rate or the "quality" of a pixel of a reference image depends on its depth information. For example, if (part of) a reference image represents a surface that is far away, then those pixels that describe that surface do not provide enough detail when the viewer zooms in or walks toward that surface. Conversely, warping every pixel of a reference image taken near an object is wasteful when the object is viewed from far away.

We characterize the reference image by a pinhole camera model using the notation adopted by McMillan [14][15]. Figure 2 illustrates the camera model. $\dot{C}$ is the center of projection. Each pixel of the reference image has coordinates (u, v) and the vectors $\vec{a}$ and $\vec{b}$ are the bases. Each pixel also contains the color information and a disparity value δ. When a pixel is projected to the 3D object space, we get a point representing the center of the projected pixel and a "stamp size." The center is computed as:

$$\dot{C} + (u\vec{a} + v\vec{b} + \vec{c}) / \delta \qquad (1)$$

and the stamp size S is calculated by:

$$S = S_X \times S_Y \qquad (2)$$
$$S_X = |\vec{a}| / \delta$$
$$S_Y = |\vec{b}| / \delta$$

To simplify our discussion, we do not consider the orientation of the object surface from which the pixel is taken. We also ignore the slight variation of stamp size at the edges of the projection plane.

An octree cell is then selected to store this pixel. The center location determines which branch of the octree to follow. The stamp size determines which level (or what size) of the octree cell should be used. The level is chosen such that the stamp size approximately matches the pixel size of the LDI in that cell.

After an octree cell has been chosen, the pixel can then be warped to the LDI of that cell. The details of the warping are described in [11]. Usually, the center of the pixel will not fall exactly on the grid of the LDI, so resampling is necessary. This is

done by splatting [20] the pixel to the neighboring grid points. In this paper we use a bilinear kernel. Four LDI pixels are updated for each pixel of a reference image. More specifically, the alpha values that result from the splatting are computed by:

$$P_X = B_X / N_X$$
$$P_Y = B_Y / N_Y$$

$$Kernel(d, s) = 1 - \frac{d}{s}$$

$$W_X = \begin{cases} Kernel(|Xi - Xc|, \frac{S_X}{P_X}), & S_X > P_X \\ Kernel(|Xi - Xc|, 1) * \frac{S_X}{P_X}, & S_X \le P_X \end{cases} \qquad (3a)$$

$$W_Y = \begin{cases} Kernel(|Yi - Yc|, \frac{S_Y}{P_X}), & S_Y > P_Y \\ Kernel(|Yi - Yc|, 1) * \frac{S_X}{P_X}, & S_Y \le P_Y \end{cases} \qquad (3b)$$

$$alpha = W_X W_Y \qquad (3)$$

where B_X and B_Y are the sizes of the LDI projection plane (which is a face of the bounding box). N_X and N_Y are the resolutions of the LDI. S_X and S_Y are as defined in equation 2. (Xc, Yc) is the center of splatting in the selected LDI and (Xi, Yi) is one of the grid points covered by the splatting. The conditions in equations 3a and 3b guarantee that the splat size will not be smaller than the LDI grid size, which represents the maximal sampling rate of the LDI.[2]

A pixel also contributes to the parent cell and all ancestor cells of the octree cell that was initially chosen. This is done by splatting the pixel to the LDIs of all the ancestor cells. The result is that the LDI of a cell contains the samples within its descendants filtered down to its resolution. Therefore, later in the rendering stage, we need not traverse the children cells if the current cell already provides enough detail.

We classify the pixels in the LDI tree into two categories: *unfiltered* and *filtered*. The unfiltered pixels are those that come from the splatting to the octree cell that was initially chosen for a reference image pixel. Those pixels that come from the splatting to the ancestor cells are classified as filtered, because they represent lower frequency components of the unfiltered pixels. Note that an unfiltered pixel may be merged with a filtered pixel during the construction of LDI tree. The merged pixel is considered as filtered because better-sampled pixels are in the LDIs of some children cells of the current octree cell.

The classification of unfiltered and filtered pixels is necessary for rendering the output images (as described in section 3.2). Imagine that a cell contains unfiltered pixels of a surface area that is only visible from one of the reference images. When the cell and its children cells are processed during the rendering, we must warp its unfiltered pixels but not its filtered pixels that are filtered from the children cells.

---

[2] It is similar to how the subpixels are prefiltered in supersampling for antialiasing.
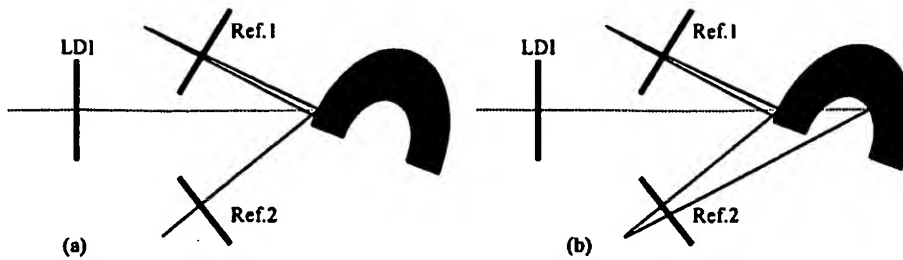
Figure 3: Illustrations of pixels that are warped to the same pixel location in an LDI. (a) Two pixels from reference image 1 and a pixel from reference image 2 are taken from the same region of a surface. Blending is used to combine their contribution to the LDI pixel. (b) One of the pixels from reference image 2 is taken from a different surface. A separate layer in the LDI is created to accommodate its contribution to the same LDI pixel.
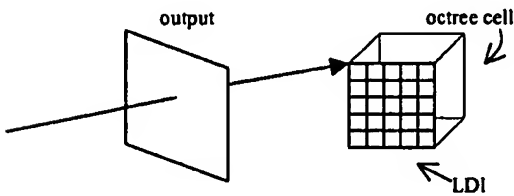

Figure 4: To estimate the range of stamp size for all pixels in the LDI, the corners of the bounding box are warped to the output image.

An LDI pixel may get contributions from many pixels of the same surface. They may be neighboring pixels in the same reference image, or pixels in different reference images that sample the same surface. The contributions from those pixels must be blended together. Figure 3a shows an example of those cases. An LDI pixel can also get contributions from many pixels of different surfaces. In those cases, we assign them to different layers of the LDI pixel. Figure 3b shows an example of those cases. To determine whether they are from the same surface or not, we check the difference in their depth value against a threshold. We select the threshold to be slightly smaller than the spacing between adjacent LDI pixels, so that the sampling rate of a surface that is perpendicular to the projection plane of the LDI can be preserved.

3.2. Rendering the Output Images

We render a new view of the scene by warping the LDIs in the octree cells to the output image. The advantage of having a hierarchical model is that we need not render every LDI in the octree. For those cells that are farther away, we can render them in less detail by using the filtered samples that are stored in the LDIs higher in the hierarchy.

To start the rendering, we traverse the octree from the top-level cell (i.e. the root). At each cell, we first perform view frustum culling, then check whether it can provide enough detail if its LDI is warped to the output image. If the current cell does not provide enough detail, then its children are traversed. An LDI is considered to provide enough detail if the pixel stamp size covers about one output pixel. Therefore the traversal of the LDI tree during the rendering will adapt to the resolution of the output image. Note that we do not calculate the pixel stamp size for each individual pixel in an LDI. Because all the pixels in the LDI of an octree cell represent samples of objects that are within its bounding box (as shown in Figure 4), we can estimate the range of stamp size for all pixels of the LDI by warping the LDI pixels that

correspond to the corners of the bounding box. The corners of the bounding box are obtained by placing the maximal and minimal possible depth at the four corner pixel locations of the LDI. We use equation 2 to compute the stamp size with the vector $\bar{a}$ and $\bar{b}$ of the output image and the disparity value δ obtained from the warping. Note that a special case exists if the new viewpoint is within the octree cell. When this happens we consider the cell as not providing enough detail and the children are traversed.

The pseudo code for the octree traversal follows:

```
Render (Octree) {
1.  If outside of view frustum,
    then return;
2.  Estimate the stamp size of the LDI
    pixels;
3.  If LDI stamp size is too large or the
    viewer is inside the bounding box then {
4.      Call Render() recursively for each
        child;
5.      Warp the unfiltered pixels in LDI to
        the Output buffer; }
6.  else {
7.      Warp both unfiltered and filtered
        pixels in LDI to the output buffer; }
}
```

Note the difference in step 5 and step 7 of the pseudo code. As mentioned in section 3.1, each LDI in the octree contains both unfiltered and filtered pixels. When we warp both the LDI in a parent cell and the LDI in a child cell, the filtered pixels in the parent cell should not contribute to the output because the unfiltered pixels in the child cell already provide better sampling for the same part of the scene.

One feature of the original LDI is that it preserves the occlusion compatible order in McMillan's 3D warping algorithm [13][14]. However this feature is compromised in the LDI tree. Although the back-to-front order can still be obtained within an LDI and across LDIs of sibling cells of the octree, we cannot obtain such order between LDIs of a parent cell and a child cell. This causes problems when unfiltered samples exist in both parent and child cells. In addition, the warped pixels are semi-transparent due to the splatting process. Therefore, we need to keep a list of pixels for each pixel location in the output buffer. We implement the output buffer as an LDI. At the end of the rendering, each list is composited to a color for display. The details of the compositing are discussed next.
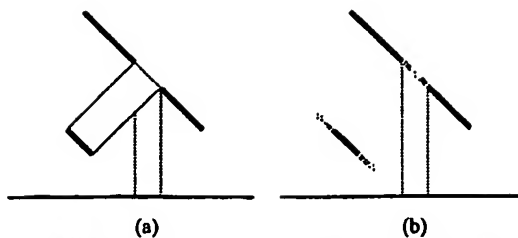
294

Figure 5: This example shows the different results of gap filling from the meshing method and the method presented in this paper. (a) The meshing method. (b) The gap filling method using filtered samples.

3.3. Compositing in the Output Buffer

Given a list of semi-transparent pixels, we sort the pixels in depth and then use alpha blending starting from the front of the sorted list. An exception is that two pixels with similar depth should be merged first and their alpha values summed together before they are alpha-blended with the other pixels. That is because they are likely to represent sampling of the same surface.

Therefore, the pixel merging is also performed in the output LDI, which is similar to the pixel merging in the LDI of the octree cell as discussed in section 3.1. The difference is that a single threshold value of depth difference does not work anymore because the pixels can come from different levels of the LDI tree. This difficulty is solved by attaching the level of octree cell where the pixel comes from to each pixel in the output LDI. The threshold value that is used for that level of octree is then used to determine whether two pixels in the output LDI should be merged.

3.4. Progressive Refinement

As discussed in section 3.2, the traversal of the LDI tree during the rendering depends on the resolution of the output image. The simplest method to create the effect of progressive refinement is to render the LDI tree to a low-resolution output image first, then increase the resolution gradually. However, this method does not utilize the coherence between the renderings of two different resolutions.

To utilize the coherence between two renderings, we can tag the octree cells that are traversed in the previous rendering and skip them in the current rendering. Note that some filtered pixels may have been warped to the output buffer if they are from the leaf nodes of the subtree traversed in the previous rendering[3]. Those pixels must also be tagged so they can be removed from the output buffer if the leaf nodes in the previous rendering become interior nodes in the current rendering.

3.5. Gap Filling

When we construct the LDI tree from many reference images, chances are we have eliminated most of the disocclusion artifacts. However, it is possible that some disocclusion artifacts still remain. We propose a two-pass algorithm that uses the filtered pixels in the LDI tree to fill in the gaps in the output image. The algorithm consists of the following steps:

1. The first pass is to render the output image from the LDI tree as discussed in section 3.2.

[3] See line 7 of the pseudo code in section 3.2.

2. A stencil (or coverage of pixels) is then built from the output image.
3. Render the output image from the LDI tree again. But in this pass, splat only the filtered pixels.
4. Use the stencil from step 2 to add the image from step 3 to the image from step 1.

The stencil from step 2 allows the filtered pixels to draw only to the gaps in the output image from step 1. This assumes that the output image would be completely filled if no disocclusion artifact occurred.[4]

Our gap filling method produces different results from the meshing method described in Mark's Post-Rendering 3D Warping [11]. Figure 5 shows an example of the gap that is caused by a front surface occluding a back surface. In the meshing method, the gaps are covered by quadrilaterals stretching between the front surface and the back surface (figure 5a). In contrast, our gap filling method splats the filtered samples from surfaces that surround the gap in the output. As shown in figure 5b, the back surfaces make more contribution to the gap than they do in the meshing method. If we do not have additional surface connectivity information in the original reference image, we believe the methods like ours that are based on the filtering of existing samples are more robust.

3.6. Analysis of Memory Requirement

Although a complete, fully expanded LDI tree may contain too many LDIs to be practical for implementation, it is worth noting that only a small subset of a complete LDI tree is used when it is constructed from reference images.

When we construct the LDI tree from reference images, we add a constant number of unfiltered LDI pixels to the octree cell chosen for each pixel of reference images. We also add $O(h)$ filtered LDI pixels to the ancestor cells, where h is the number of ancestors. That means the amount of memory taken by the LDI tree grows in the same order as the amount taken by the original reference images, only if h is bounded.

We can further assume that h is bounded because the maximal height of the LDI tree exists. Let L be the longest side of bounding box of the scene, N be the resolution of an LDI, d be the smallest feature in the scene the human eyes can discern at a minimum distance, and H be the maximal height of the LDI tree. Then we have:

$$H = \left\lceil \log_2 \frac{L}{N \times d} \right\rceil$$

Although we do not include the memory overhead for maintaining the octree, we also do not include the possible saving in memory when pixels are merged in the LDIs. The experimental results will be presented later in this paper to show that amount of memory indeed grows at a slower rate than the number of reference images.

3.7. Rendering Time

An advantage that image-based rendering has over traditional polygon-based rendering is that the rendering time does not grow with the complexity of the scene. That advantage is still preserved in the rendering from the LDI tree, even though more layers of LDIs must be rendered. Let us consider the worst case in which we need to render every pixel in the LDI tree. As discussed

[4] See previous footnote [1] for special cases such as the windows in the video and figure 11.

previously, the number of pixels grows in the same order as the original reference images. Therefore the time complexity of rendering from the LDI tree is of the same order as warping all reference images in the worst case. Because larger cells are used for farther objects, the worst case rarely happens and usually much fewer pixels in the LDI tree are rendered. The experimental results are presented in the next section.

4. RESULTS

We implemented the LDI tree on a Silicon Graphics Onyx2 with 16 gigabytes of main memory. The machine has 32 250 MHz MIPS R10000 processors but we did not exploit its parallel processing capability in our implementation.

We tested our program with a model of the interior of Palladio's II Redentore in Venice [16]. The reference images are generated by ray tracing using the Rayshade program [5]. Each reference image has 512x512 pixels and 90-degree field of view. Figure 6 shows one of the reference images.

In synthesized scenes, an LDI can be generated directly by ray tracing [19]. We do not include it in our framework because it does not apply to the reference images acquired from non-synthesized scenes, such as the depth images that are acquired by a laser range finder.

Figure 7 shows the top view of the bounding boxes of the LDI tree after two of the reference images are processed. Each cell has an LDI of 64x64 resolution. The left face of each cell is also the projection plane of its LDI. Note that the cells near the center of projection of a reference image have more levels of subdivision. Figure 8 shows a new view rendered from the LDI tree. We disabled the gap filling to let the disocclusion artifacts appear in blue background color. Figure 8 has severe disocclusion artifacts because only four reference images from the same viewpoint are used. Figures 9 and 10 show the same view but with 12 and 36 reference images (from 3 and 9 viewpoints) respectively. Figure 11 is generated from the same LDI tree as figure 10 but with the gap filling enabled.

The memory usage of the LDI trees is shown in chart 1. The first reference image consumes about 30 Mbytes (MB) of memory. About 15 MB is the overhead of the octree. The resampling and filtering (described in section 3.1) generates about 5 LDI pixels for each input pixel. As more reference images are added, the growth of the memory size slows. The last 60 images add less than 1 MB per image in average. Note that the growth of the memory size does not stop completely. That is because more detail near each new viewpoint is still being added to the LDI Tree.

Chart 2 shows the rendering time for various numbers of reference images. Each line represents the rendering times along the path for a given number of reference images. The priority in our experiment is the correctness. Therefore little optimization and hardware acceleration were used to speed up the rendering. For example, the splatting operation is implemented completely in software simulation.

Chart 3 shows the growth of the (averaged) rendering time when the number of reference images increases. It shows that the rendering time grows even slower than the size of memory because some unnecessary details added from additional reference images are not processed during the rendering.

5. CONCLUSION AND FUTURE WORK

Using multiple reference images in 3D image warping has been a challenging problem. This paper describes the LDI tree, which
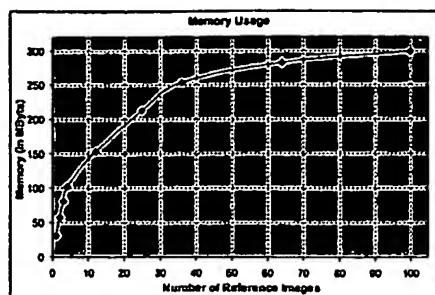

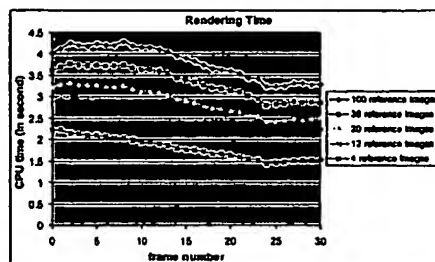Chart 1: The memory usage of LDI trees.
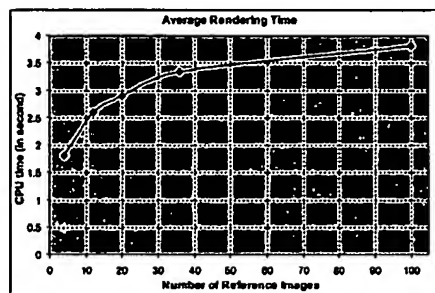

Chart 2: The rendering time.

Chart 3: The average rendering time per frame.

combines multiple reference images into a hierarchical representation and preserves their sampling rate of the scene. The LDI tree allows the efficient extraction of the best available samples for any view and uses filtered samples in the hierarchy to reduce the rendering time. The filtered samples also enable the gap filling method presented in section 3.5.

We have assumed that each pixel of reference images provides only the color and depth information. No surface normal or orientation information has been considered. A direction for future work is to incorporate the surface orientation into our framework, for use in the splatting and the calculation of stamp size.

When a surface is sampled in multiple reference images, we should be able to get better sampling of the surface than what we can get from any single image. How to explore this type of cross-image supersampling is another direction of future work.

Like the original LDI, pixels that fall into the same pixel location and have similar depth values are merged together. That is based on the assumption that the surface is diffuse and little view-dependent variance can occur. How to extract view-dependent properties of the surface is yet another direction for future work.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] C. H. Chien, Y. B. Sim and J. K. Aggarwal. Generation of Volume/Surface Octree from Range Data. The Computer Society Conference on Computer Vision and Pattern Recognition, pages 254-60, June 1988.

[2] C. I. Connolly. Cumulative Generation of Octree Models from Range Data. Proceedings, Intl' Conf. Robotics, pages 25-32, March 1984.

[3] Brian Curless and Marc Levoy. A Volumetric Method for Building Complex Models from Range Images. In Proceedings of SIGGRAPH 1996, pages 303-312.

[4] Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski and Michael F. Cohen. The Lumigraph. In Proceedings of SIGGRAPH 1996, pages 43-54.

[5] Craig Kolb. Rayshade. http://www-graphics.stanford.edu/~cek/rayshade/.

[6] David Laur and Pat Hanrahan. Hierarchical Splatting: A Progressive Refinement Algorithm for Volume Rendering. Computer Graphics (SIGGRAPH 91 Conference Proceedings), volume 25, pages 285-288.

[7] Marc Levoy and Pat Hanrahan. Light Field Rendering. In Proceedings of SIGGRAPH 1996, pages 31-42.

[8] A. Li and G. Crebbin. Octree Encoding of Objects from Range Images. Pattern Recognition, 27(5):727-739, May 1994.

[9] Dani Lischinski and Ari Rappoport. Image-Based Rendering for Non-Diffuse Synthetic Scenes. Rendering Techniques '98 (Proc. 9th Eurographics Workshop on Rendering).

[10] Robert W. Marcato Jr. Optimizing an Inverse Warper. Master's of Engineering Thesis, Massachusetts Institute of Technology, 1998.

[11] William R. Mark, Leonard McMillan and Gary Bishop. Post-Rendering 3D Warping. Proceedings of the 1997 Symposium on Interactive 3D Graphics, pages 7-16.

[12] Nelson Max. Hierarchical Rendering of Trees from Precomputed Multi-Layer Z-Buffers. Rendering Techniques '96 (Proc. 7th Eurographics Workshop on Rendering), pages 165-174.

[13] Leonard McMillan. A List-Priority Rendering Algorithm for Redisplaying Projected Surfaces. Technical Report 95-005, University of North Carolina at Chapel Hill, 1995.

[14] Leonard McMillan and Gary Bishop. Plenoptic Modeling. In Proceedings of SIGGRAPH 1995, pages 39-46.

[15] Leonard McMillan. An Image-Based Approach to Three-Dimensional Computer Graphics. Ph.D. Dissertation. Technical Report 97-013, University of North Carolina at Chapel Hill. 1997.

[16] Nathan O'Brien. Rayshade - Il Redentore. http://www.fbe.unsw.edu.au/exhibits/rayshade/church/

[17] Gernot Schaufler and Wolfgang Stürzlinger. A Three-Dimensional Image Cache for Virtual Reality. In Proceedings of Eurographics '96, pages 227-236. August 1996.

[18] Jonathan Shade, Dani Lischinski, David H. Salesin, Tony DeRose and John Snyder. Hierarchical Image Caching for Accelerated Walkthrough of Complex Environments. In Proceedings of SIGGRAPH 1996, pages 75-82.

[19] Jonathan Shade, Steven Gortler, Li-wei He and Richard Szeliski. Layered Depth Images. In Proceedings of SIGGRAPH 1998, pages 231-242.

[20] Lee Westover. SPLATTING: A Parallel, Feed-Forward Volume Rendering Algorithm. Ph.D. Dissertation. Technical Report 91-029, University of North Carolina at Chapel Hill. 1991.
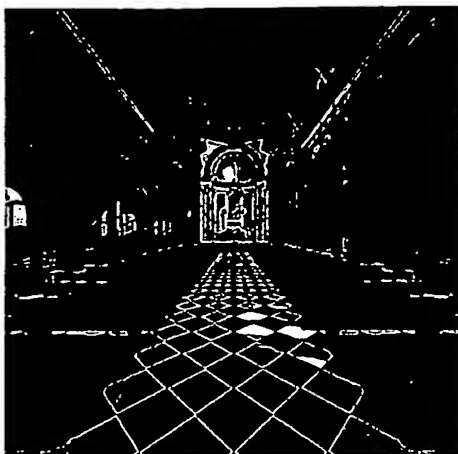
Figure 6: One of the reference images.

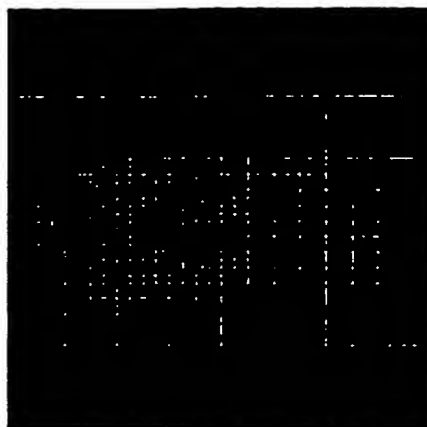Figure 7: Top view of the octree cells after combining two reference images.

Figure 8: A new view generated from four reference images (at the same position).

Figure 9: A new view generated from 12 reference images (at three different positions).

Figure 10: A new view generated from 36 reference images (at 9 different positions).

Figure 11: A new view generated from 36 reference images. Gap filling is enabled.

298

PORTAL

USPTO

Octrees for faster isosurface generation

Authors   Jane Wilhelms   Univ. of California, Santa Cruz
Allen Van Gelder Univ. of California, Santa Cruz

Additional Information: abstract references citings index terms review collaborative colleagues peer to peer

Tools and Actions: Find similar Articles   Review this Article

Save this Article to a Binder   Display Formats: BibTex EndNote ACM Ref

DOI Bookmark: Use this link to bookmark this Article: http://doi.acm.org/10.1145/130881.130882
What is a DOI?

↑ ABSTRACT

The large size of many volume data sets often prevents visualization algorithms from providing interactive rendering. The use of hierarchical data structures can ameliorate this problem by storing summary information to prevent useless exploration of regions of little or no current interest within the volume. This paper discusses research into the use of the octree hierarchical data structure when the regions of current interest can vary during the application, and are not known a priori. Octrees are well suited to the six-sided cell structure of many volumes. A new space-efficient design is introduced for octree representations of volumes whose resolutions are not conveniently a power of two; octrees following this design are called branch-on-need octrees (BONOs). Also, a caching method is described that essentially passes information between octree neighbors whose visitation times may be quite different, then discards it when its useful life is over. Using the application of octrees to isosurface generation as a focus, space and time comparisons for octree-based versus more traditional "marching" methods are presented.

↑ REFERENCES

Note: OCR errors may be found in this Reference List extracted from the full text article. ACM has opted to expose the complete List rather than only correct and linked references.

1   ARTZY, E., FRIEDER, G., AND HERMAN, G. The theory, design, implementation, and evaluation of a three-dimensional surface detection algorithm. Comput. Graph. Image Process. 15, 1 (Jan. 1981), 1-24.

2   Jon Louis Bentley, Multidimensional binary search trees used for associative searching, Communications of the ACM, v.18 n.9, p.509-517, Sept. 1975

3   J. Bloomenthal, Polygonization of implicit surfaces, Computer Aided Geometric Design, v.5 n.4,

Octrees for Faster Isosurface Generation

JANE WILHELMS and ALLEN VAN GELDER
University of California, Santa Cruz

The large size of many volume data sets often prevents visualization algorithms from providing interactive rendering. The use of hierarchical data structures can ameliorate this problem by storing summary information to prevent useless exploration of regions of little or no *current* interest within the volume. This paper discusses research into the use of the *octree* hierarchical data structure when the regions of current interest can vary during the application, and are not known *a priori*. Octrees are well suited to the six-sided cell structure of many volumes.

A new space-efficient design is introduced for octree representations of volumes whose resolutions are not conveniently a power of two; octrees following this design are called *branch-on-need octrees* (BONOs). Also, a caching method is described that essentially passes information between octree neighbors whose visitation times may be quite different, then discards it when its useful life is over.

Using the application of octrees to isosurface generation as a focus, space and time comparisons for octree-based versus more traditional "marching" methods are presented.

Categories and Subject Descriptors: E.1 [Data]: Data Structures—*trees*; I.3.3 [Computer Graphics]: Picture/Image Generation—*display algorithms*; I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—*curve, surface, solid, and object representations*; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—*visible line/surface algorithms*

General Terms: Algorithms, Performance

Additional Key Words and Phrases: Hierarchical spatial enumeration, isosurface extraction, octree, scientific visualization

---

1. INTRODUCTION

Interactive visualization is of major importance to scientific users, but the sheer size of volume data sets can tax the resources of computer workstations. Intelligent use of data structures and traversal methods can make a

---

significant difference in algorithm performance. In particular, the use of hierarchical data structures to summarize volume information can prevent useless traversal of regions of little interest. However, the storage and traversal of hierarchical data structures themselves can add to the resource consumption of the algorithm, both in terms of time and space.

We are exploring the advantages and disadvantages of hierarchical data structures for visualization. In particular, we have explored the use of *octrees* in conjunction with a cell-oriented isosurface generation algorithm [13, 26, 27].

Octrees are particularly appropriate for representing sample data volumes common to scientific visualization, where the data points often define a spatial decomposition into hexahedral, space-filling, nonoverlapping regions. Use of octrees for controlling volume traversal is appropriate whether regions are regular hexahedra (cubes, rectangular parallelopipids), as is common in medical imaging, or the irregular, warped hexahedra (curvilinear decompositions) that are common in computational fluid dynamics.

A volume whose maximum resolution is between 2^{k-1} and 2^k can be represented by an octree of depth k. This paper discusses the use of summary information at each node for the entire subvolume beneath it, making it possible to explore the volume contents without examining every data point. For isosurface generation the summary information consists of the maximum and minimum values of data within each node's region.

1.1 Background and Prior Work

Octrees, like quadtrees, are hierarchical data structures based on decomposition of space [14–17, 20–22]. Quadtrees are two-dimensional decompositions that had their beginnings in the hierarchical representation of digital image data and spatial decomposition for hidden surface elimination [11, 19, 24]. In quadtrees, space is recursively subdivided into four subregions, hence the name "quad". Octrees are three-dimensional extensions of quadtrees, where space is recursively subdivided into eight subvolumes, and the root of the octree refers to the entire volume [15–17, 22]. In the normal case, each coordinate direction is divided in two, giving a "lower" half space and an "upper" half space. The effect of all three divisions is to create octants.

Octrees have been used to represent three-dimensional objects [10, 28]. Octrees have also been used just to represent the spatial relationship of geometrical objects, making it relatively simple to accomplish such operations as locating neighbors [18] and to traverse the volume from front to back for hidden surface removal [4, 25].

In many octree applications, including those mentioned so far, the octree is used to represent some boolean property of the points in the volume, or some property for which most of the points take on a null value that is specified *a priori*. In image-processing terminology, a point in the volume is "black" (in the object), or "white" (uninteresting). Here we briefly review some storage optimizations that have been developed for such cases, and discuss why they do not carry over to the applications we have, in which the volume data can assume many values (none of which may be "uninteresting" *a priori*).

When the property is boolean, only one bit per octree node is needed. Levoy described a straightforward implementation for abstracting the (boolean) property of *nontransparency* from medical image data as part of volume rendering [12]. Initially, his method rounds the volume resolution up to $2^d \times 2^d \times 2^d$, and assigns eight data points to each node in the lowest level of the octree. It represents every node at the same level of the octree in a long bit-vector (1 = "black"), where 1 denotes that *some* child has value 1, or at the lowest level, that some data point is nontransparent among the eight covered by the octree node. All octree information is located by address calculations; no pointers are needed. The storage overhead is acceptable, well under 20% of the original volume data in practice.

An alternative strategy is to prune lower portions of the octree when their values can be inferred from an ancestor [29]. One method is to define an internal node as "white" or "black" if all of its descendants are of that color, in which case no storage is allocated to the descendants; this process is called *condensation*. Otherwise the node is gray and has 8 explicit children. (For static nonboolean properties, only white nodes can be condensed.) How many octree nodes are needed depends on the original data. Because of the irregular shapes possible in such octrees, the structure must be represented explicitly, with pointers being the usual choice. Eight pointers per node use up storage quickly, so this implementation is workable only when the object can be represented with relatively few black and white nodes. However, it is possible to reduce the storage requirement to one pointer per node if all eight children of a node are allocated contiguously.

Linear octrees were introduced by Gargantini as a way to improve on the storage requirements of condensed, pointer-based octrees [6]. Related linear structures were used by others [16, 23]. Essentially, each "black" node in the condensed octree is assigned a key that encodes the path in the octree from the root to that node (see Section 4.1). Gray and white nodes are not allocated any storage, and the keys of the black nodes are stored in sorted order in one array (hence the name "linear"). Whether a linear octree requires more or less storage space than a bit-vector octree depends on the coherence of the boolean property being represented.

Glassner describes an implementation related to *linear octrees*, but with several innovations [8]. He uses a hash table instead of a sorted array to speed up node location by key. His ray tracing application requires storage of gray nodes, so he uses a slightly different key and allocates all 8 children of a node contiguously, so they can all be accessed under one key entry.

Bloomenthal also uses an octree to organize nonboolean data for implicit surface modeling [3]. A closed-form function is defined over the volume and evaluated by adaptive sampling. A piecewise polygonal representation is derived from the octree. The octree only pertains to the current isovalue, or threshold value, so this application also falls into the category of those whose data has a frequently occurring null value.

We are concerned here with the use of octrees to organize nonboolean data, where the points of interest cannot be determined *a priori*; that is, there is no frequently occurring null value. The reason that the condensation methods

just discussed are not applicable in this context soon becomes evident: condensation occurs only when all children of a node have the same value, an event that may *never* occur in volumetric data such as density fields. We are not dealing with an object, or small set of objects, which occupies a possibly small portion of the volume, but rather a function that is defined throughout the volume. As we shall show in Section 3, without the benefits of large scale condensation, obvious octree designs can easily lead to prohibitive storage overhead.

Globus has independently investigated the use of an octree for isosurface generation [9]. His work is compared with ours in more detail in Sections 3.4 and 6.2. Briefly, he solved the storage problems by stopping the octree construction at a higher level, allowing an octree node to cover as many as 32 data points.

An alternative tree data structure for 3-dimensional data is the 3-*d tree*, which is a special case of the *k*-d tree for *k*-dimensional data. A *k*-d tree is really a binary tree in which each node divides in some coordinate direction [2, 21]. The choice of direction can depend on the region "covered" by the node. By choosing to split the root in the *z* direction, to split the nodes at depth one in the *y* direction, those at depth two in the *x* direction, and repeating that cycle, we can simulate an octree with a 3-d tree. 3-d trees might offer advantages similar to octrees in some situations.

1.2 Summary of Results

Unlike earlier octree applications, where certain regions of the volume could be classified as uninteresting *a priori*, we studied the use of octrees to organize volume information when all of the volume is potentially interesting. In Section 4 we present an octree design that has proven to be efficient in time with acceptable storage requirements, which we call a *branch-on-need octree* (BONO). Appendices A and B provide a technical supplement to this section. In Section 3 we show that more obvious alternate designs will have prohibitive storage requirements in most practical cases, where the resolutions of the volume are not precisely $2^d \times 2^d \times 2^d$, as is usually conveniently assumed.

An important time-saver in isosurface generation is the reuse of computed information on cell edges that intersect the isosurface; each such edge is incident on four cells, so the computation can be used four times if it can be saved and located. A normal coordinatewise (marching) traversal of the volume permits a straightforward caching strategy with arrays [13]. However, the octree traversal order complicates storage-efficient caching considerably. We solved the problem with a hash table, as described in Section 5 and Appendix C. A key feature of the solution is that we can tell when a hash table entry has been retrieved for the last time, and delete it, making room for later entries.

Section 6 presents our experimental results, which compare the performance of an octree-based isosurface generation program with the more standard, nonhierarchical methods, such as marching cubes [13], and its variants [26]. In applying octrees to isosurface generation, it is important to

remember that the only part of the processing that we are addressing is the detection and bypassing of trivial cells: those that do not intersect the current isosurface. Isosurface patches are calculated in significant cells with the same subroutines as used by the marching traversal, so their time cost is unchanged. Therefore, it is somewhat surprising that our experiments demonstrated speedups by factors of 2 and 3 in some cases, even when octree creation time is included.

Because one application often generates many isosurfaces from the same data, the speedup on the surface extraction phase alone is often more significant to the user. We observed speedups in the range of 1.6 to 11.

We develop a performance model based on the experimental data to predict the time requirements of isosurface generation with our implementations of both octree traversal and marching traversal.

Rounding out the paper, Section 2 reviews polygon-based isosurface generation methods, and describes our adaptations of previous methods to take advantage of an octree, and Section 7 draws some conclusions and suggests future directions for the research.

2. OCTREES IN CONJUNCTION WITH ISOSURFACE GENERATION

A common approach to visualization is to extract a geometrical representation of a surface of constant threshold value (the *isosurface*) from sampled volume data. Graphics workstations are deft at handling such geometrical representations efficiently, offering the ability to render hundreds of thousands of polygons per second. For large volumes with complex surfaces, however, generation of the geometric representation may take many minutes.

A popular method for isosurface generation is to imagine the volume as consisting of cells whose corners are the sample values [3, 5, 13, 26, 27]. Each cell is examined one by one for the presence of an isosurface, which is detected when at least one corner value is above and another below the threshold value. If the isosurface intersects the cell, intersection points along the cell edges are calculated and become the vertices of polygons representing the portion of the isosurface within that cell. Lorensen and Cline introduced a table-lookup method to speed polygon generation [13].

Generally, isosurfaces intersect a small subset of the cells within a volume. However, most of the useful work of the algorithm occurs within those cells that do intersect the isosurface. The relative costs of traversal versus cellular computation are extremely variable, depending upon the total size of the volume, the number of cells including the isosurface, and the size of the computer memory. Previous research indicated that between 30% and 70% of the time spent in isosurface generation was spent examining empty cells [26]. This provided impetus for the study of the octree traversal methods described here.

2.1 Two Contrasting Approaches

We explored two approaches to isosurface generation: the first is a typical *marching* method [13, 26] and the second is the *octree-traversal* method

introduced here. Both methods read the volume data into an array, begin with a setup phase, then continue with a surface-finding phase for each threshold furnished by the user.

The marching method has a minimal setup phase; for the user's convenience in selecting thresholds, it finds the maximum and minimum data values. Each surface-finding phase visits all cells of the volume, normally by varying coordinate values in a triple "for" loop. As each cell that intersects the isosurface is encountered, the necessary polygons to represent the portion of the isosurface within the cell are generated. There is no attempt to "trace" the surface into neighboring cells. To find the isosurface for a new threshold value the whole phase is repeated; there is no carry-over information.

During its setup phase, the octree method creates an octree that contains at each node the maximum and minimum data values found in that node's subtree. The lowest level of the octree represents eight cells, and contains a pointer into the data array to the sample value having the minimum (x, y, z) value of any of the 27 samples defining these eight cells. The volume data is stored in an ordinary 3-D array, rather than octree traversal order, to simplify the location of neighboring data points. In contrast to the marching method, the setup phase does a substantial amount of work, and determining maxima and minima are an essential part of the setup, not merely a user convenience.

In surface-finding phases, the octree is traversed with a particular threshold, only exploring those branches that contain part of the isosurface; any node whose maximum is below the threshold or whose minimum is above it is exited without traversing its children. When a leaf node that contains isosurface is visited, each of the (normally eight) cells that it "covers" are visited, and polygons are generated.

Both methods use a table lookup for polygon generation. The basic idea is due to Lorensen and Cline [13]; refinements to handle "ambiguous" cells were described by Wilhelms and Van Gelder [26]. The table contains 256 entries, referring to the 256 combinations of positive and negative (relative to threshold) values that can occur for an eight-cornered cell. Each table entry describes which cell edges contain intersections and how they should be joined to produce the polygons representing the isosurface. A second table is used to treat ambiguous cases; each ambiguous case in the first table contains a "pointer" to the relevant section of the second table.

3. SPACE REQUIREMENTS OF PREVIOUS OCTREE DESIGNS

The space requirements of an octree can be a serious issue in the design of a system that will process large data volumes. In this section we examine space requirements of previous designs; in Section 4 we describe a more space-efficient design. First, we review octree basics and introduce some terminology.

3.1 Octree Basics

Octrees are tree structures of degree eight. It is convenient to number the children from zero to seven; their numbers, written in binary, encode which

subregion of the parent they "cover". We shall use the zyx convention. If the z bit is 1, the child covers an octant that is "upper in z"; if it is 0, the child covers an octant that is "lower in z". The y and x bits are similarly interpreted. We shall write child numbers in binary to facilitate this interpretation. (An alternate notation is back/front for z, south/north for y, and west/east for x.)

Traversal of an octree is accomplished by recursively visiting a node and traversing its children in order. Notice that all children of a fixed node that are "lower in z" are visited before all children that are "upper in z"; among those that are in the same z division, the ones that are "lower in y" are visited first, etc.

A *full octree* is one in which each node has exactly eight children; however, this is possible only if the volume's resolution is the same power of two in each dimension, e.g., $4 \times 4 \times 4$ (64) samples, $8 \times 8 \times 8$ (512) samples, $64 \times 64 \times 64$ (262,144) samples, etc. Full octrees offer the best ratio of the number of nodes to data points. For a volume with a resolution of s in each direction where s is a power of 2,

$$nodes = \sum_{i=0}^{\log_2 s - 1} 8^i = \frac{s^3 - 1}{7}.$$

As described in Section 3.3, the regularity of the full octree data structure permits it to be implemented without storing explicit addressing information in the octree node; we call this a *pointerless* octree design. An alternative is to use an octree design in which nodes are allocated space only if they "cover" a region that is actually within the volume. This makes the location of nodes less predictable. Traditionally, each node contains addressing information necessary to determine the location of its children. We call this design a *pointer* octree.

3.2 A Running Example

Previous treatments of octrees have made the simplifying assumption that the resolutions of the volume are precisely $2^d \times 2^d \times 2^d$ for some integer d. However, power-of-two volumes are not the norm; moreover, volumes often vary widely in resolution among the three dimensions. In this case, storage of a full octree can be extremely wasteful because many nodes correspond to regions not actually within the volume.

To explore the impact on previous designs when the power-of-two assumption does not hold, we shall consider an example at some length. To make the discussion concrete, let us assume that each data value requires the same space as a pointer or index, and call this a "word". Usually a "word" is 32 bits. For our application, isosurface generation, each octree node must store two words (*maximum* and *minimum*), plus whatever structural bookkeeping is required.

For our running example, consider a data volume whose x, y, and z resolutions are $320 \times 320 \times 40$, for a total of 4,096,000 data points. An octree

for this volume will have nine levels. This is the size of one of our CT-scan volumes for which computational experience is presented; see Section 6.

3.3 Pointerless Full Octree Design

Suppose we naively set up a "full" octree over this 320 × 320 × 40 volume; that is, every node in the octree has exactly eight children, and each leaf node refers to eight data points (which may or may not be within the actual volume). One motivation for using a full octree is that the nodes can be stored in an array T in such a way that parent and child pointers are not required. In analogy with the heap-sort strategy in one dimension and the pyramid strategy in two dimensions, the root is in $T[0]$, the children of the node occupying $T[k]$ are found in locations $T[8k + 1] \cdots T[8k + 8]$, and all nodes at the same level are contiguous within the array. (Thus we can think of $T[9] \cdots T[72]$ as a subarray containing all the nodes at depth two, etc.)

Unfortunately a full octree for the example volume requires

$$1 + 2^3 + 4^3 + \cdots + 256^3 = 19{,}173{,}961$$

nodes of two words each, or almost 40 million words. (As discussed earlier, in applications requiring only one *bit* per node the full octree fits in slightly over 500,000 words, which is quite acceptable.) This overhead in this example, almost 4 times as many nodes as data points and nearly eight times as much space as the original volume, is almost certainly not acceptable. Most of the space is wasted, but is not easy to eliminate because "real" nodes are scattered throughout it.

3.4 Traditional Design of Pointer Octrees

An alternative is to build the octree in the more traditional way, with pointers or indices (subscripts) to a node's children within the node record. Nodes are only created if they "cover" some portion of the data. Leaf nodes should require only one pointer, which is to data, because the neighbors of that point can be located. Although a naive design would specify a pointer for each child, giving at least eight pointers per internal node, with some care, all of a node's children can be allocated contiguously, so that one child pointer (or index) suffices for internal nodes as well. In this design, each octree node occupies three words. (Some designs might include a fourth word for a parent pointer.)

Let us see how this might work, following the traditional and intuitive *even-subdivision* strategy, which divides each node's range from the top down in each coordinate direction as evenly as possible. (Ranges of 1 or 2 are not divided.) Consider an octree in which the three resolutions are not equal. Because nodes branch in each dimension from the top down until no more subdivisions are required, the even-subdivision strategy will result in 8-way branching at the top when all dimensions subdivide, 4-way branching in the middle when two dimensions subdivide, and (effectively) binary subtrees at the bottom when only the largest dimension continues to subdivide. As most nodes are at the greater depths, this means the least efficient nodes are the

Node depth	Number of nodes	Region covered
0	1	320x320x40
1	8	160x160x20
2	64	80x80x10
3	512	40x40x5
4	4,096	20x20x3 or 20x20x2
5	24,576	10x10x2 or 10x10x1
6	98,304	5x5x2 or 5x5x1
7	393,216	3x3x2, 3x2x2, 2x2x2 or 3x3x1, 3x2x1, or 2x2x1
8	786,432	2x2x2, 2x2x1, or 2x1x1
total	1,307,209	

Fig. 1. An even-subdivision octree covering a 320 × 320 × 40 data volume.

most numerous; consequently, the ratio of nodes to data points can be quite high.

The outcome for our example is shown in Figure 1. Observe that a node has null children whenever one or more of its coverage resolutions is 1 or 2. (Strategies to detect when this occurs are not difficult, and are similar to those employed in our actual implementation, as discussed later.) When all coverage resolutions are two or less, the node is a leaf, and it points to data.

As Figure 1 shows, an octree designed by this strategy for our example consists of about 1,300,000 nodes, and nearly four million words. This octree requires almost as much memory as the original volume—better than a full pointerless octree, but still a serious overhead. Even if a pointerless strategy were devised for this octree, the use of even subdivision will produce an octree whose size is about 2/3 of the original volume. The ratio of octree nodes to data points is 0.3191, a significant degradation when compared to the optimum of 0.1428. This observation motivated our search for an improvement.

Globus reports a variation of the even-subdivision strategy that addresses the storage space issue [9]. Primarily, a coarser granularity is accepted in that an octree node that covers less than 32 cells is not further subdivided. Also, nodes that cover small but very oblong regions are divided four or eight times in the longest dimension, and not divided in one or both of the shorter dimensions. On our example Globus' strategy yields an octree of 201,289 nodes; each leaf node turns out to cover 25 cells in a 5 × 5 × 1 pattern. The number of words needed per node depends on implementation choices that were not reported; various trade-offs between time and space are possible. Comparison with his timing results appears in Section 6.2.

4. A SPACE-EFFICIENT OCTREE DESIGN

This section describes the octree design we adopt and compares the space requirements with the even-subdivision method. Essentially, we regard the

octree as conceptually full, but avoid allocating space for empty subtrees. Note that the even-subdivision strategy divides the volume, from the top down, whenever it can. The approach we describe now, in some sense, *delays* subdivision until absolutely necessary. Therefore we call it the *branch-on-need* strategy, and we call the resulting data structure a *branch-on-need octree* (BONO for short). The presentation here is from a top-down point of view, because the procedures work top-down to facilitate storage allocation. An alternative bottom-up view is discussed in Section 4.2.

4.1 Branch-on-Need Octrees

We can associate with each node a conceptual region and an actual region, as illustrated in Figure 2 with our running example, which is a 320 × 320 × 40 volume. Recall that in our terminology, "001 child" means the "lower z, lower y, upper x" child. The three bits of the child code represent motion in the z, y, and x directions, respectively, when read left to right. Since all of the "upper z" children of the root have empty actual regions, no space is allocated for them. Therefore the root has only four actual children, and we say that it "branches" in the x and y directions, but not in the z direction.

A further element of the BONO strategy is that the "lower" subdivision in each branching direction always covers the largest possible exact power of two (yielding a range of the form $2^k - 1$). A two-dimensional analog contrasting the even-subdivision and branch-on-need strategies is shown in Figure 4 on a 5 × 6 array.

Definition 4.1. We define the *range* in each of the x, y, and z directions as the difference between the upper and lower limits of the actual region. The *range vector* is the triple of x, y, and z ranges.

An interesting pattern emerges if we look at a node's range vector in binary. In our example, for the root we have:

Direction	Range in Binary	(Decimal)
x	100111111	(319)
y	100111111	(319)
z	000100111	(39)

The directions that branch are precisely those that have a 1 bit in the leftmost position, when all ranges are written with the same number of bits. The number of bits equals the height of the node in the octree, with leaves considered height 1 (and data considered height 0). To obtain the range of the "upper" children in a direction that branches, simply remove that leftmost 1 bit. The range of the corresponding "lower" children is a bit string of 1's, one shorter than the root's original range. Following this rule, the ranges of the 001 child of the root are:

Direction	Range in Binary	(Decimal)
x	00111111	(63)
y	11111111	(255)
z	00100111	(39)

Node	Conceptual Region			Actual Region		
	x	y	z	x	y	z
root	0–511	0–511	0–511	0–319	0–319	0–39
000 child	0–255	0–255	0–255	0–255	0–255	0–39
001 child	256–511	0–255	0–255	256–319	0–255	0–39
010 child	0–255	256–511	0–255	0–255	256–319	0–39
011 child	256–511	256–511	0–255	256–319	256–319	0–39
100 child	0–255	0–255	256-511	0–255	0–255	*empty*
. . .	. . .	. . .	. . .	. . .	. . .	*empty*

Fig. 2. Conceptual and actual regions for octree nodes over a 320 × 320 × 40 volume.

We see immediately that this node branches in the y direction, but not the x or z directions. Both of its children will have the following range configuration:

Direction	Range in Binary	(Decimal)
x	0111111	(63)
y	1111111	(127)
z	0100111	(39)

The bit patterns of the ranges also allow us to quickly discover how many (actual) nodes the octree has at each depth. This information is vital for the allocation of storage. For example, to see how many nodes are at depth 4, we take the 4 leftmost bits of each range of the root:

Direction	Range in Binary
x	1001|11111
y	1001|11111
z	0001|00111

This gives ranges of 9, 9, and 1. Add 1 to each and multiply, giving 10 • 10 • 2 = 200 nodes at depth 4. To see why this works, just imagine that we *began* with a data volume of resolutions 10 × 10 × 2, and built an octree over it. Then the root's ranges would be 9, 9, and 1.

Figure 3 shows the number of nodes produced by this strategy on our 320 × 320 × 40 example (4,096,000 data points). In contrast to previous schemes, the 585,439 nodes are far fewer than the number of data points, and the ratio is virtually the optimum one of 1 node per 7 data points for a full octree. The space, about 1,750,000 words, is well under 50% of that occupied by the data volume. This behavior is typical, and not an artifact of the resolutions chosen for the example. As shown rigorously in Appendix A, when all resolutions are at least 32, the ratio of octree nodes to data points never exceeds 0.1615, which is not far from the optimum of 0.1428.

Using the above observations, we can efficiently allocate precisely the correct amount of space for an octree, and determine which directions any

Node depth	Number of nodes	Region covered
0	1	320x320x40
1	4	256x256x40 or 256x64x40
		or 64x256x40 or 64x64x40
2	9	128x128x40 or 128x64x40
		or 64x128x40 or 64x64x40
3	25	64x64x40
4	200	32x32x32 or 32x32x8
5	1,200	16x16x16 or 16x16x8
6	8,000	8x8x8
7	64,000	4x4x4
8	512,000	2x2x2
total	585,439	

Fig. 3.   Nodes by depth for a branch-on-need octree (BONO) covering a 320 × 320 × 40 volume.

given node branches. Our implementation precomputes this information and stores in each octree node a 3-bit code to tell which directions branch, and an index to its "leftmost" (000) child. The actual children of the node are contiguous in the array in lexicographic order. That is, if all eight children are actual, their order is 000, 001, 010, 011, 100, 101, 110, 111. We assume the volume has less than 2^{28} data points and pack the code and index into one 32-bit word.

In fact, all nodes at a given depth are contiguous and appear in what we call *shuffled zyx* order. For example, at depth 4, a node's origin in the data volume is the triple

Direction	Origin in Binary
x	$x_9 x_8 x_7 x_6 000000$
y	$y_9 y_8 y_7 y_6 000000$
z	$z_9 z_8 z_7 z_6 000000$

where x_9 is 0 for the lower-in-x children of the root, and is 1 for the upper-in-x children of the root, etc.

The node's *shuffled zyx* key is

$$z_9 y_9 x_9 z_8 y_8 x_8 z_7 y_7 x_7 z_6 y_6 x_6.$$

Interpreting the bits of this key in groups of three gives the path in the octree to this node.

Appendix B describes how to calculate the location of a node without using pointers (or indices) from its key and the range vector of the root of the octree. Although this can be done in time proportional to the length of the key, it is still fairly expensive, so we chose to incur the space overhead of one index per node to speed traversal of the octree. As the discussion showed, this

space overhead is about 15% of the basic data volume. Furthermore, it can be greatly reduced by use of pointerless nodes on the leaf level only.

Our isosurface application does not require the ability to locate an arbitrary node in the octree, but this is a necessary operation for many other algorithms [6, 7, 8, 23]. For example, Gargantini uses (essentially) the same shuffled *zyx* key to specify a node's place in the octree [6]. However, her *linear octree* explicitly stores the key with the node, and stores the nodes sorted in key order; a binary search is employed to locate a node in memory by its key. Glassner uses a slightly different key and a hash table [8]. In contrast, we do not store the key at all, but can calculate the node's location in memory from the key.

4.2 Comparison of Branch-on-Need and Even-Subdivision Strategies

The branch-on-need strategy can also be viewed as a *bottom-up* one, as compared to the top-down even-subdivision strategy described earlier. Data points are grouped from the bottom up in the most efficient manner. For volumes with different resolutions in the three dimensions, this places the 8-way branching section of the hierarchy (or 8-way collapsing if one thinks from the bottom up) at the bottom of the tree; 4-way branching occurs where one dimension has been clustered as much as possible, and an (effectively) binary tree occurs where only one dimension still requires clustering. Thus, the most efficient reductions take place at the bottom of the tree where there are the most nodes. In general, this strategy is space-wise far superior to the even-subdivision approach. Figure 4 demonstrates a two-dimensional analog of this phenomenon on a 5×6 array; in three dimensions and with larger resolutions the difference is much more pronounced. The two methods are identical and produce a full octree when the volume resolutions are precisely $2^d \times 2^d \times 2^d$.

.Two factors influence the ratio of nodes to data points for branch-on-need octrees: the size of each resolution (resolutions of a power of two are best, and of a power of two plus one are worst); and the relative size of the resolutions (cubical volumes are best). The size of each resolution is a more significant influence, but neither effect seriously degrades the ratio. We show this in an intuitive manner below; Appendix A gives the formal proof. Consider volumes all of whose resolutions are at least 32. (This bound is chosen as a reasonable minimum resolution. The true worst case ratio can be produced by having two resolutions of 1. The tree is then a binary tree, which is uninteresting.)

First, consider a volume that contains $(s + 1) \times (s + 1) \times (s + 1)$ data points, where s is a power of two. A full octree can be built over $s \times s \times s$ data points, which will contain $\frac{1}{7}(s^3 - 1)$ nodes. Three quadtrees can be built over the remaining nodes covering three faces; each one will contain $\frac{1}{3}(s^2 - 1)$ nodes. Three binary trees can be built along the three edges which still include nodes not considered, each adding approximately s nodes. The necessity of nodes with only two or four children on the deepest levels of the tree produces the worst case, which is approximately $\frac{1}{7}s^3 + s^2$ nodes actually needed, compared with the optimum of about $\frac{1}{7}(s + 1)^3$. While not a major degradation, the ratio of actual to optimum is about $(1 + 4/s)$.
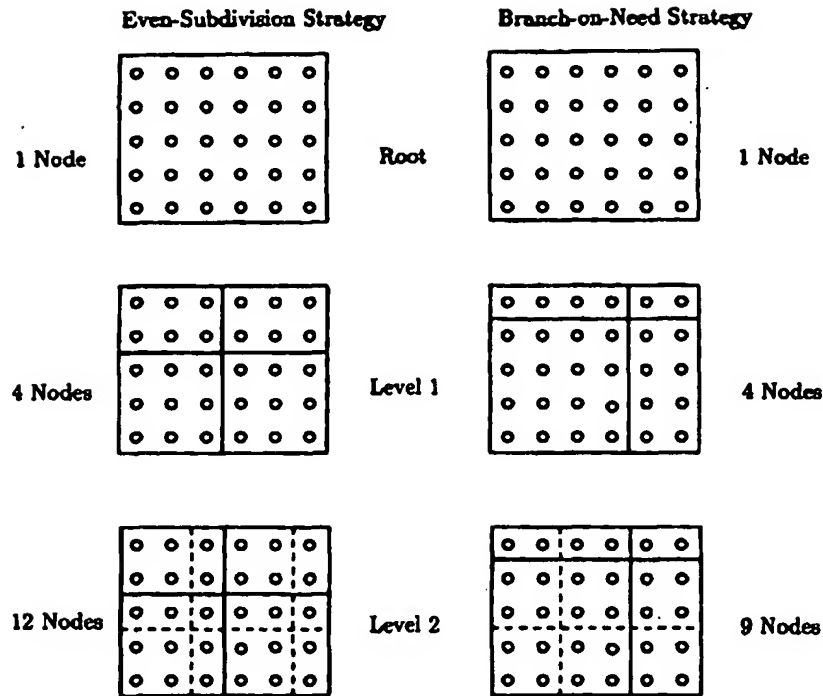
Even-Subdivision Strategy   **Branch-on-Need Strategy**

Fig. 4.   Comparison of designs in two dimensions on a 5 × 6 array.

For larger volumes, up to a resolution of $2s \times 2s \times 2s$, the "binary" and "quad" nodes in the above described tree can be given new children, making them more "efficient", and the ratio of nodes to data points will never be worse than for the above case. For each power of two over 32, the ratio improves slightly, approaching $\frac{1}{7}$. Thus, the worst case among cubical volumes is one with resolutions $33 \times 33 \times 33$. This volume requires 5803 nodes for 35937 data points, a ratio of .1615.

What about the effect of one resolution being much larger than the other? Consider the effect of a volume of resolutions $s \times s \times t$, where s is a power of two and at least 32, as before, and $t \gg s$. An octree of height $\log_2 s$ can be built over the volume. At the top of this octree, two resolutions cannot undergo further divisions and the other resolution is s. A binary tree of height $\log t - \log_2 s$ can be built on top of the octree. The octree contains approximately $\frac{1}{7}s^2 t$ nodes and the binary tree approximately t/s nodes. The effect of the binary tree is relatively negligible; the ratio of nodes to data points remains near $\frac{1}{7}$.

The branch-on-need strategy normally produces a tree shaped quite differently from the even-subdivision version. The even-subdivision strategy will partition the volume into more equal parts. Using a one-dimensional exam-

ple, if there are 65 data points, the even-subdivision strategy will divide at the top into one subtree covering 33 data points and another covering 32. The branch-on-need strategy will divide into one region covering 64 data points and the other covering 1 point.

In rare cases, it may be necessary to visit more nodes in a BONO than in an even-subdivision octree to process a particular section of the volume. In this one-dimensional example, if the two highest-indexed points were the only ones of interest, the branch-on-need tree would require visitation of 13 nodes because these two points are on separate subtrees of the root. The even-subdivision tree would require only six or seven. However, there are counterexamples, as when the two points are in the center of the volume, where the even-subdivision tree requires more visitation because the nodes are on separate subtrees of the root while in the BONO tree they are not. Furthermore, as traversal has a modest cost (see Section 6), this whole issue is not a major consideration.

5. RECALLING PREVIOUSLY COMPUTED INTERSECTION POINTS

Each vertex is generally contained in four neighboring polygons, and each vertex requires six floating point numbers: three representing location, and three a normal vector required for rendering. Calculating vertex information, particularly the normal vector, is quite expensive, and a substantial time savings is realized by reusing the results. Storage is also saved by representing polygons by indices into a vertex array, saving over twice the space normally required for geometric representations.

A straightforward array method for saving this information, as described by Lorensen and Cline [13], is used in the marching method. However, it is not suitable for octrees because the octree traversal does not visit the nodes in row-major order. It is possible to devise a savings method designed particularly for use with octrees, but a hash table appeared to provide a simple and general solution to the problem. Wyvill et al. also used a hash table in their implementation [27]. Technicalities of our hash table are given in Appendix C.

The main observation needed for storage efficiency here is that it is possible to identify the last visit of an edge, and remove its information from the hash table, freeing the storage for later use. Because of traversal order, the three edges adjacent to the "origin" of a cell will never be visited again. This allowed us to use hash tables which are much smaller than would be necessary to store all significant edges simultaneously. Artzy et al. used a related storage optimization in their traversal of an implicit binary spanning tree; although their data structure was a set of linked lists, the removal of "marked nodes" that would never be visited again was critical [1].

6. EXPERIMENTAL RESULTS

Tests were run on six sets of data. In all cases, use of octree traversal was faster than the marching approach. This was true whether the time for octree creation was included or whether only the actual surface-finding traversal

Table I.   Characteristics of Data Volumes

Data File	Resolution	Number of Cells	Octree Size	Threshold	% Cells with Surface
Blunt Fin	40x32x32	31,117	5,855	1.0	12.97%
Protein	64x64x64	226,981	37,449	0.1	0.82%
Enzyme	97x97x116	998,468	160,383	36.5	9.22%
Dolphin	320x320x40	3,718,093	585,439	120.2	2.87%
NMR Brain	256x256x109	6,784,954	1,032,229	500.5	7.04%
CT Head	256x256x113	7,040,990	1,070,373	150.5	4.28%

times were compared. The justifications for comparing the two methods on the surface-finding phases only are two: first, the octree can be precomputed and stored for reuse; and, second, it is possible to use the same octree for multiple thresholds.

6.1 Description of Experimental Data

Table I describes the data. The *blunt fin* (C. M. Hung and P. G. Buning, NASA Ames Research Center) is a curvilinear volume generated using computational fluid dynamics and extracting a surface from the density field. The superoxide dismutase *enzyme* (D. M. McRee, Scripps Clinic) and the high-potential iron *protein* (L. Noodleman and D. Case, Scripps Clinic) are molecular volumes from the volume data set distributed by the University of North Carolina. The *dolphin* (T. Cranford, UC Santa Cruz) is a threshold from a CT-scan of a dolphin head, using only central slices. The *MR-brain* (Siemens Medical Systems) and the *CT-head* (North Carolina Memorial Hospital) are also scans from the UNC dataset. Figure 5 shows the images of some of these surfaces generated on a Silicon Graphics Iris.

6.2 Experimental Timing Results

Table II summarizes the costs of the two methods on the seven sets of data. Runs were made on a Sun Sparcstation 1 with eight megabytes of memory. Tests on a 16 megabyte machine have produced similar relative timing results. Note that, in general, octree traversal shows increasingly better relative performance as the data files get larger. This is to be expected, as the fraction of cells containing isosurface tends to decrease as volume size increases.

The *octree-creation* time includes allocating memory for the octree, recursively traversing it downwards establishing pointers, and accumulating maximum and minimum information on the return to the root node. The *surface-finding* time involves, for the marching version, traversal of all cells, and, for the octree version, traversal of the regions of the octree and volume as dictated by summary information in the nodes. The *total generation* time is the sum of these for the octree method, and is the same as the surface-finding
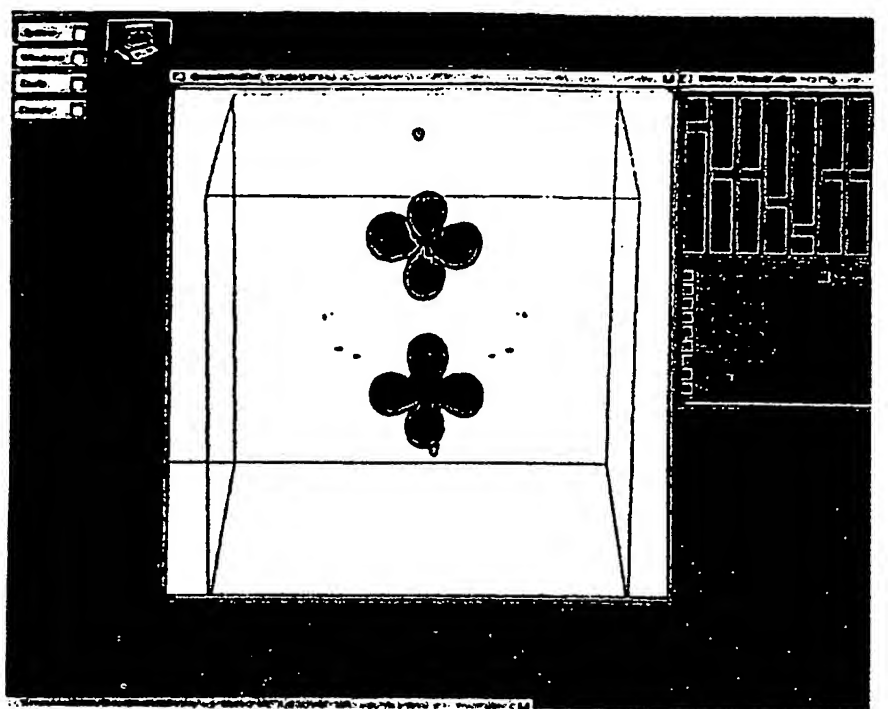
(a)

Fig. 5.   Selection of isosurfaces analyzed. (a) Iron protein wave function threshold .1; (b) Enzyme electron density map threshold 36.5; (c) NMR-scan of head threshold 500.5.

time for the marching method. The time to display the resultant polygons would be the same for both methods, and is not included. The time to read in the data and do minor preliminary initialization is not included in the statistics either, because it is the same for both methods. It is worth mentioning that for very large files, this cost is approximately equal to the octree creation time.

In the best experimental case, the protein, the octree method improved surface-finding speeds by a factor of 11. For this volume, relatively few cells have surfaces and these are concentrated in small regions of the volume. For the other volumes, surface-finding using the octree took between a quarter and two thirds the time of the standard method.

The actual octree traversal times were insignificant. For example, the MR Brain volume took under 4.5 seconds for actual octree traversal, compared to almost 300 seconds for the surface-finding phase.

Globus reports experiments on the same blunt fin data set that we used [9]. He does not report precise thresholds used, but the threshold we used should be most comparable to his maximum case. There we both found that the
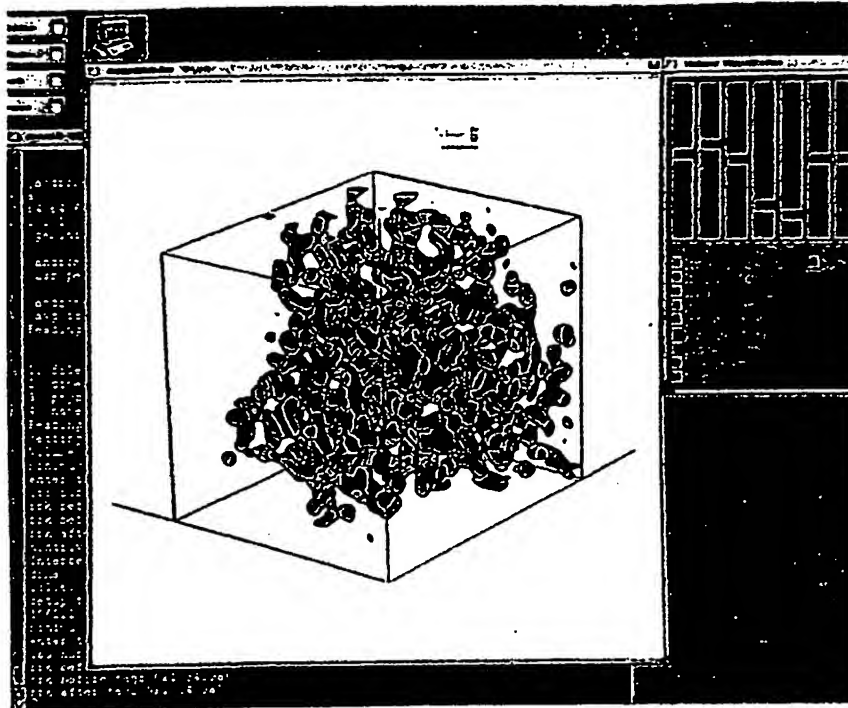
(b)
Fig. 5—*Continued*

octree produces a substantial time reduction in the surface-finding phase: he
reports about 4.7 versus 8.9, or 53% to compare with our 64%. However, he
experiences a somewhat larger relative cost for building the octree: 1.78/8.9,
or 20% of the "marching" time. We found it to be .36/3.00, or 12%. This
difference may be due to a more complicated subdivision strategy (see Section
3.4). In any event we both observe a benefit even when only one surface is
extracted, with increasing dividends as the octree is reused for subsequent
surfaces. Globus, like us, found substantially greater speedups in other cases,
including some by a factor of 9.9.

6.3 Performance Models of Surface Finding

We estimated a linear function of the total number of cells and the number of
cells intersecting the isosurface to explain the running times observed in our
experiments. Specifically, the running time is modelled as $time = At + Bs$,
where t is the total number of cells actually visited and s is the number with
isosurface. The estimates of constants A and B are shown below. The value
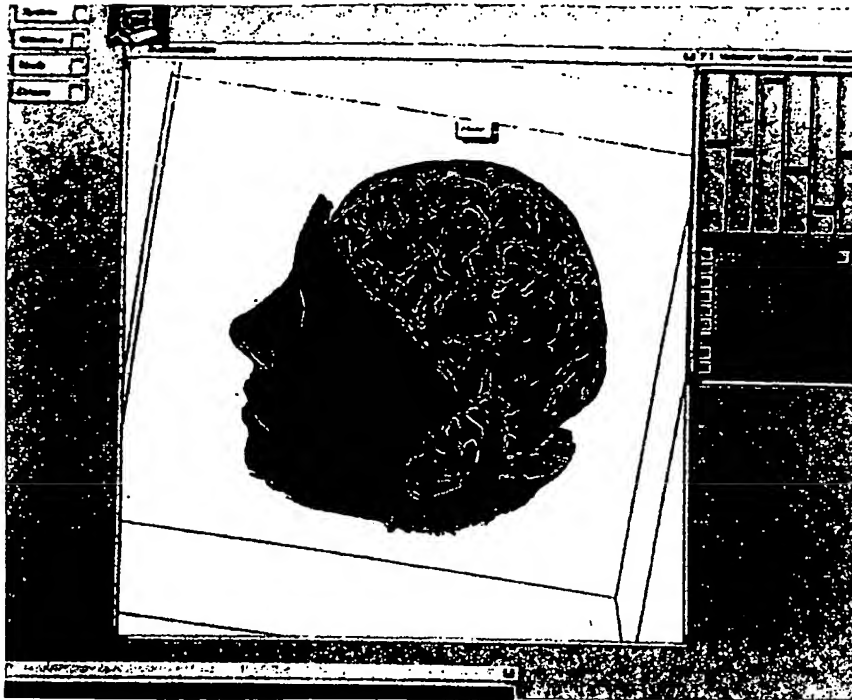
(c)

Fig. 5—*Continued*

of *A* for the octree includes the overhead of traversing the internal nodes.

Surface-finding times in μ-secs.	march	octree
per cell visited (*A*)	54	72
additional time if cell intersects isosurface (*B*)	431	431

We do not have enough data to achieve statistical significance; these values were informally estimated, and fit the larger runs better than the smaller ones.

We observed that the octree method very consistently visited about twice as many cells as had isosurface, because the lowest internal node indicates whether an isosurface may be present in any of up to eight cells. Thus, $t = 2s$ for the octree method, to a good approximation. Of course, t equals the whole volume for marching methods. From this crude model we can estimate a *ratio* of surface-finding times of the two methods as a function of f, the *fraction* of cells that intersect the isosurface. Let r be the ratio of marching time to octree time for the surface-finding phase. Let t now be the total number of

Table II.   Comparative Processing Times for Isosurface Generation, in CPU Seconds

	March Traversal	Octree Traversal	% Octree/March
Blunt Fin			
Octree Creation		0.36	
Surface Finding	3.00	1.90	64%
Total Extraction	3.00	2.26	75%
Protein			
Octree Creation		2.50	
Surface Finding	11.1	0.95	9%
Total Extraction	11.1	3.45	31%
Enzyme			
Octree Creation		11.9	
Surface Finding	75.2	43.0	57%
Total Extraction	75.2	54.9	73%
Dolphin			
Octree Creation		56.5	
Surface Finding	224.9	60.7	27%
Total Extraction	224.9	117.2	52%
MR Brain			
Octree Creation		109.6	
Surface Finding	556.5	282.2	51%
Total Extraction	556.5	391.8	70%
CT Head			
Octree Creation		114.1	
Surface Finding	464.1	167.1	36%
Total Extraction	464.1	281.2	61%

cells in the volume; thus $s = ft$. We have

$$r = \frac{54t + 431ft}{72(2ft) + 431ft} = .75 + \frac{.094}{f}.$$

Whenever f is less than about .37, r exceeds 1, and we anticipate that the octree method will outperform the marching methods *in the surface-finding phase*. We have yet to encounter a situation where the percentage of cells with surface is anywhere near this.

The other side of the coin is that the octree method requires significant setup time, other than reading in the data. For this time factor, we estimated $C = 16$ μ-seconds per cell in volume. (The comparative value for the marching method is just the time per cell to get the maximum and minimum of the volume, and is optional; it is about 5 μ-seconds.)

As remarked before, the setup time can often be amortized over several surface-finding phases. However, for a run consisting of setup and finding one isosurface, we get the ratio r_1 marching time to octree time for the run:

$$r_1 = \frac{5t + 54t + 431ft}{16t + 72(2ft) + 431ft} = .75 + \frac{.082}{f + .028}.$$

Whenever f is less than about .30, r_1 exceeds 1, and we anticipate that the octree method will outperform the marching methods for single isosurface runs (including both setup and surface-finding phases). For f less than about .037, we calculate that octrees are better by a factor of 2.

6.4 Miscellaneous Remarks

Use of an octree traversal without a saving strategy to reuse previous intersection-point-related computations obviated its speed advantages. Preliminary results showed the octree traversal was about equivalent to the traditional marching method when the octree did not save those computations. Similarly, removal of the saving strategy from the traditional method significantly slows down that algorithm. Therefore, use a saving strategy.

Another interesting side note is that traversal times on machines with relatively small memories can be highly dependent on traversal order. Assume the x dimension varies fastest in the array that stores the data volume, followed by y, and then z. We inadvertently found ourselves at one time traversing the volume in the order z-y-x and found that it could take many times as long as an x-y-z traversal, due to the time taken by page faults. Our particular tests were on an 8 megabyte machine, which is surely "relatively small memory" by current standards in graphics. But small memory is relative: Moving to a machine with greater capacity soon leads us to work with larger volumes. While there is no reason to use z-y-x order for polygon generation methods, algorithms such as direct volume rendering normally access the volume front to back, and hence direction can make a significant difference. In such cases, storing the data in octree order to equalize traversal costs might be preferred.

Even very small data sets stored in ascii take a long time to read and convert. Ascii is most convenient for portability, but should be converted to floats or integers, as appropriate, for repeated use.

7. CONCLUSIONS AND FUTURE RESEARCH

Our studies showed that octrees can yield substantial improvements in performance for isosurface generation on data sets produced by current technologies. These improvements will be even more significant on larger data sets. However, several technical problems had to be solved to realize the benefits.

(1) A new octree implementation made the storage overhead acceptable.

(2) A careful caching method enabled us to reuse results of earlier computations, then discard them when they would not be referenced again, freeing the storage for others.

Numerous topics remain to be explored concerning octrees in scientific visualization. The use of octrees on irregular, nonhexahedral grids, as are often produced in finite-element analysis, requires further research. Other applications of octrees should also present new, interesting, technical problems.

APPENDIX A: RATIO OF BRANCH-ON-NEED OCTREE NODES TO DATA POINTS

In this appendix we show that the BONO design creates a number of nodes that is less than 1/6 the number of data points covered, for volumes all of whose dimensions are at least 32. The lower bound of 32 was chosen as a bound that should be satisfied in practice. The same analysis can be repeated with looser restrictions to get slightly weaker bounds. Recall that 1/7 is the best possible case, and is easily achieved by all methods when the volume is $2^d \times 2^d \times 2^d$.

To simplify the formulas, we use the nomenclature introduced in Definition 4.1 that the *range* in each coordinate direction is one less than the number of points (or resolution) in that direction. Thus a $32 \times 32 \times 48$ volume has the range vector (31, 31, 47).

Let (x_0, y_0, z_0) denote the range vector for the data. Let (x_h, y_h, z_h) denote the range vector for the octree nodes at height h, where the leaves of the octree are at height 1. As discussed earlier, from the way the octree is formed,

$$x_h = \lfloor x_0/2^h \rfloor$$
$$y_h = \lfloor y_0/2^h \rfloor$$
$$z_h = \lfloor z_0/2^h \rfloor.$$

That is, h rightmost bits are dropped off each range coordinate at height zero.

Throughout this discussion let d be the height of the root, and let $1 \le h \le d$. We are concerned with bounding the ratio

$$R(x_0, y_0, z_0) = \sum_{h=1}^{d} \frac{(x_h + 1)(y_h + 1)(z_h + 1)}{(x_0 + 1)(y_0 + 1)(z_0 + 1)}$$

for ranges $x_0 \ge 31$, $y_0 \ge 31$, $z_0 \ge 31$. Thus the minimum *volume* to which our bound applies is $32 \times 32 \times 32$.

To find the point where $R(x_0, y_0, z_0)$ achieves its maximum, we use the fact that R must be maximized in any coordinate direction when the other two coordinates are held constant. The analysis is complicated by the fact that there are numerous local maxima. To simplify the formulas that follow, we abbreviate:

$$K_h = \frac{(y_h + 1)(z_h + 1)}{(x_0 + 1)(y_0 + 1)(z_0 + 1)}.$$

First we consider points such that x_0 is a power of 2, and compare $R(2^m, y_0, z_0)$ for different values of m.

LEMMA A.1. $R(2^m, y_0, z_0)$ is a decreasing function of m, for $1 \leq m < d$ and fixed y_0 and z_0.

PROOF. $R(2^{m+1}, y_0, z_0) - R(2^m, y_0, z_0)$ can be written as

$$\sum_{h=1}^{m} \left(\frac{(2^{m+1-h} + 1)}{(2^{m+1} + 1)} - \frac{(2^{m-h} + 1)}{(2^m + 1)} \right) K_h$$

$$+ \left(\frac{2}{(2^{m+1} + 1)} - \frac{1}{(2^m + 1)} \right) K_{m+1}$$

$$+ \sum_{h=m+2}^{d} \left(\frac{1}{(2^{m+1} + 1)} - \frac{1}{(2^m + 1)} \right) K_h .$$

The second sum is clearly negative (or zero if $m + 1 = d$). The other terms can be simplified to

$$\sum_{h=1}^{m} \left(\frac{(2^{m-h} - 2^m)}{(2^{m+1} + 1)(2^m + 1)} \right) K_h$$

$$+ \left(\frac{1}{(2^{m+1} + 1)(2^m + 1)} \right) K_{m+1} .$$

Using the facts that $m \geq 1$ and $K_{m+1} \leq K_m$, we see that the mth term of the sum is at least as negative as the term under the sum is positive. Since all terms of the sum are negative, the lemma is proved.   □

Next we show that a power of two dominates all the values up to the next power of two.

LEMMA A.2. $R(2^m, y_0, z_0) > R(2^m + k, y_0, z_0)$ for $1 \leq k < 2^m$, $1 \leq m < d$, and fixed y_0 and z_0.

PROOF. $R(2^m + k, y_0, z_0) - R(2^m, y_0, z_0)$ can be written as

$$\sum_{h=1}^{m} \left(\frac{(2^{m-h} + \lfloor k/2^h \rfloor + 1)}{(2^m + k + 1)} - \frac{(2^{m-h} + 1)}{(2^m + 1)} \right) K_h$$

$$+ \sum_{h=m+1}^{d} \left(\frac{1}{(2^m + k + 1)} - \frac{1}{(2^m + 1)} \right) K_h .$$

The second sum is clearly negative. The first sum can be simplified to

$$\sum_{h=1}^{m} \left(\frac{(-k2^{m-h} + 2^m \lfloor k2^{-h} \rfloor - k + \lfloor k2^{-h} \rfloor)}{(2^m + k + 1)(2^m + 1)} \right) K_h .$$

Since all terms of the sum are negative or zero, the lemma is proved.   □

Now we can establish a worst-case bound on the ratio of BONO nodes to data points:

THEOREM A.3.   *For $x_0 \geq 31$, $y_0 \geq 31$, and $z_0 \geq 31$, the maximum value of $R(x_0, y_0, z_0)$ occurs at $(32, 32, 32)$, corresponding to a $33 \times 33 \times 33$ data volume.*

PROOF.   Lemmas A.1 and A.2 show that no $x_0 > 32$ can produce the maximum, and direct calculation shows $R(31, y_0, z_0) < R(32, y_0, z_0)$, so R must be maximized at $x_0 = 32$. The same arguments apply to y_0 and z_0.   □

The ratio of BONO nodes to data points for a $33 \times 33 \times 33$ volume is

$$R(32, 32, 32) = \frac{17^3 + 9^3 + 5^3 + 3^3 + 2^3 + 1}{33^3}$$

$$= .1615 < \tfrac{1}{6}$$

which is the worst case for volumes all of whose dimensions are at least 32.

APPENDIX B: LOCATION OF NODES FROM KEYS

Here we sketch the calculation of a node's location, given its depth in the octree, its shuffled *zyx* key, and the range vector of the octree. The underlying idea is that nodes at the same depth preceding the desired node are in the subtree of a smaller sibling of some ancestor of the desired node. The relationship to Gargantini's *linear octrees* is discussed near the end of Section 4.1.

Assume the key is a list of octal digits describing the path from the root to the desired node. Let function head return the first element of such a list, and let tail return the list of all elements except the first. Also, let childRange(child, range) return the range vector of a child, given the child number (an octal digit), and range, the range vector of the parent, as described in Section 4.

The function offset(key, range, depth) shown in Figure 6 returns the number of actual nodes at depth depth whose shuffled *zyx* keys are lexicographically less than key in an octree whose range vector is range. This number gives the offset of the desired node from the beginning of the subarray of nodes at that depth.

The depth of recursion equals the depth of the desired node, which is proportional to the length of the key. The work at each level is bounded by a constant, as the for loop body executes at most seven times. Thus the function takes time that is linear in key length. Nevertheless, a substantial time penalty would be incurred if this calculation were used instead of a pointer, to save space.

APPENDIX C: HASH TABLE DESIGN DETAILS

This appendix describes some details of the hash table used to store edge-related calculations for later use in the octree traversal, as discussed in Section 5. Other designs are certainly workable, but this can provide a starting point for other implementors.

```
offset(key, range, depth)
    {
    count = 0;
    if (depth > 0)
        {
        ancestor = head(key);
        for (sibling = 0; sibling < ancestor; sibling++)
            {
            s = childRange(sibling, range);
            count += (s.x + 1) * (s.y + 1) * (s.z + 1);
            }
        count += offset(tail(key), childRange(ancestor, range), depth-1);
        }
    return count;
    }
```

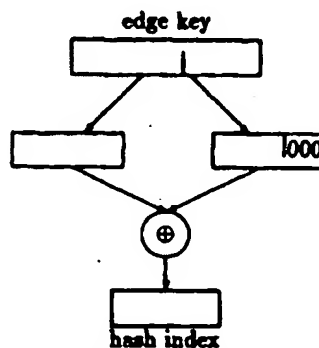Fig. 6. Procedure to calculate node location from shuffled *zyx* key.

Fig. 7. The hash function, where "⊕" denotes "exclusive or".

Each edge in the volume is given a unique key. Assume the edge is from (x, y, z) to a point one greater in some coordinate direction. Then the offset of (x, y, z) in the volume array (viewed now as one-dimensional) is the basis for the key. The "direction-code" assigned is 1 for x, 2 for y, or 3 for z. The key is then

$$4 * \text{offset} + \text{direction-code}$$

The "null" key is 0.

As an example, consider the edge from $(3, 3, 3)$ to $(3, 4, 3)$ in a 320 × 320 × 40 volume. The offset of $(3, 3, 3)$ is 308,163, and this is a y edge, so its key is 1,232,654.

Our experience showed that a good size for the hash table is eight times the square root of the number of data points in the volume, rounded up to a power of two. We wanted a fast hash function that would distribute edges in the same cell. We settled on one shown diagrammatically in Figure 7; its

mathematical form is

$$(K/F) \oplus 8(K \bmod F)$$

where $\oplus$ denotes *exclusive* or, K is the key and F, the "fold point", is 1/8 the size of the hash table. (This function relies on K being less than $8F^2$.) Upon collision, we rehashed by adding 1 mod table size.

We tracked utilization of the hash table, and found that it normally got about one quarter full, and averaged about one collision per operation.

REFERENCES

1. Artzy, E., Frieder, G., and Herman, G. The theory, design, implementation, and evaluation of a three-dimensional surface detection algorithm. *Comput. Graph. Image Process. 15,* 1 (Jan. 1981), 1–24.
2. Bentley, J. L. Multidimensional binary search trees used for associative searching. *Commun. ACM 18,* 9 (1975), 509–517.
3. Bloomenthal, J. Polygonization of implicit surfaces. *Comput. Aided Geom. Des. 5,* 4 (Nov. 1988), 341–355.
4. Doctor, L. J., and Torborg, J. G. Display techniques for octree-encoded objects. *IEEE Comput. Graph. Appl. 1,* 3 (July 1981), 29–38.
5. Gallagher, R. S., and Nactegaal, J. C. An efficient 3-D visualization technique for finite element models. *Comput. Graph. 23,* 3 (Aug. 1989), 185–194.
6. Gargantini, I. Linear octrees for fast processing of three-dimensional objects. *Comput. Graph. Image Process. 20,* 4 (Dec. 1982), 365–374.
7. Gargantini, I., Walsh, T. R., and Wu, O. L. Viewing transformation of voxel-based objects via linear octrees. *IEEE Comput. Graph. Image Process. 6,* 10 (Oct. 1986), 12–20.
8. Glassner, A. S. Space subdivision for fast ray tracing. *IEEE Comput. Graph. Appl. 4,* 10 (Oct. 1984), 15–22.
9. Globus, A. Octree optimization. In *Symposium on Electronic Imaging Science and Technology* (San Jose, Calif., Feb. 1991), SPIE/SPSE.
10. Jackins, C. L., and Tanimoto, S. Oct-trees and their use in representing 3D objects. *Comput. Graph. Image Process. 14,* 3 (Nov. 1980), 249–270.
11. Klinger, A., and Dyer, C. R. Experiments on picture representation using regular decomposition. *Comput. Graph. Image Process. 5,* 1 (Mar. 1976), 68–105.
12. Levoy, M. Efficient ray tracing of volume data. *ACM Trans. Graph. 9,* 3 (July 1990), 245–261.
13. Lorensen, W. E., and Cline, H. E. Marching cubes: A high resolution 3D surface construction algorithm. *Comput. Graph. 21,* 4 (July 1987), 163–169.

14. MAO, X., KUNII, T., FUJISHIRO, I., AND NOMOA, T. Hierarchical representations of 2D/3D gray-scale images and their 2D/3D two-way conversion. *IEEE Comput. Graph. Appl. 7*, 12 (Dec. 1987), 37-44.

15. MEAGHER, D. J. Octree encoding: A new technique for the representation, manipulation, and display of arbitrary three-dimensional objects by computer. Tech. Rep. IPL-TR-80-111. Image Processing Laboratory, Rensselaer Polytechnic Institute, Troy, N.Y., Oct. 1980.

16. MEAGHER, D. J. Geometric modeling using octree encoding. *Comput. Graph. Image Process. 19*, 2 (June 1982), 129 147.

17. MEAGHER, D. J. Interactive solids processing for medical analysis and planning. In *Proceedings NCGA 5th Annual Conference* (Dallas, Tex., May 1984), pp. 96 106.

18. MOORE, M., AND WILHELMS, J. Collision detection and response for computer animation. *Comput. Graph. (ACM Siggraph Proceedings) 22*, 4 (Aug. 1988), 289 298.

19. SAMET. H. The quadtree and related hierarchical data structures. *ACM Comput. Surv. 16*, 2 (June 1984), 186 260.

20. SAMET. H. *Applications of Spatial Data Structures.* Addison-Wesley, Reading, Mass., 1990.

21. SAMET. H. *The Design and Analysis of Spatial Data Structures.* Addison-Wesley Reading, Mass., 1990.

22. SRIHARI, S. N. Representation of three-dimensional digital images. *ACM Comput. Surv. 13*, 4 (Dec. 1981), 399 424.

23. TAMMINEN, M., AND SAMET, H. Efficient octree conversion by connectivity labeling. *Comput. Graph. (ACM Siggraph Proceedings) 18*, 3 (July 1984), 43-51.

24. WARNOCK, J. E. The hidden line problem and the use of halftone display. In *Proceedings of the Second University of Illinois Conference on Computer Graphics, Pertinent Concepts in Computer Graphics*, M. Faiman and J. Nievergelt, Eds. (Mar. 1969), pp. 154-163.

25. WATT, A. *Fundamentals of Three-Dimensional Computer Graphics.* Addison-Wesley Reading, Mass., 1st ed., 1989.

26. WILHELMS, J., AND VAN GELDER, A. Topological ambiguities in isosurface generation. Tech. Rep. UCSC-CRL-90-14. CIS Board, Univ. of California, Santa Cruz, Dec. 1990. Extended abstract in ACM *Comput. Graph. 24*, 5 (Dec. 1990), 79 86.

27. WYVILL, G., MCPHEETERS, C., AND WYVILL, B. Data structures for soft objects. *The Visual Comput. 2*, 4 (Aug. 1986), 227-234.

28. YAMAGUCHI, K., KUNII, T. L., AND FUJIMURA, K. Octree-related data structures and algorithms. *IEEE Comput. Graph. Appl. 4*, 1 (Jan. 1984), 53-59.

29. YAU, M. M., AND SRIHARI, S. N. A hierarchical data structure for multi-dimensional digital imaging. *Commun. ACM 26*, 7 (July 1983), 504-515.